

## VOICE OUTPUT UNIT AND NAVIGATION SYSTEM

### Background of the Invention

The present invention relates to a voice output unit which converts a text-based document into voice and outputs the voice thus converted, and a navigation system.

As a conventional voice output unit, there is a technical art as described in the Japanese Patent 10 Laid-Open Publication No. 2002-108378, for example.

This voice output unit aims at changing pitch or speed of the voice, when a text-based document is converted into voice, according to a hometown of a person who created the document, so that a sense of realism can be given to a listener.

### Summary of the Invention

However, in the conventional art, when this voice output unit functions as a navigation system, for example, it is not possible to recognize a hometown and the like of a document creator of any type of the document, such as road guidance and an E-mail obtained via the Internet. Consequently, the voice is outputted with a same intonation, speed and so on. Under this situation, if the road guidance comes into 25 while the listener is listening to the E-mail, there is a problem that he or she may miss the road guidance.

Focusing attention on the foregoing problem of the conventional art, the present invention is directed to providing a voice output unit and a navigation system which gives a sense of realism to a listener, even if there are plural types of documents other than narrative or the like in which a hometown or the like of a document creator can be recognized, and which also features that even when the document is switched to a different type of document, it can be easily perceived by the listener.

In order to achieve the above object, the voice output unit of the present invention comprises, a voice signal synthesizer for generating a voice signal from said text-based document, an output means which outputs as a voice said voice signal generated in said voice signal synthesizer, a grasping means which grasps contents or a length of said text-based document, and a synthesizing controller for allowing said voice signal synthesizer to generate said voice signal in which a tone quality including at least intonation is changed, according to the contents or the length of said text-based document grasped in said grasping means, when said voice signal synthesizer generates said voice signal from said text-based document.

In order to achieve the above object, the navigation system of the present invention comprises,

a voice signal synthesizer for generating a voice signal from said text-based document,

5 an output means which outputs as a voice said voice signal generated by said voice signal synthesizer,

a grasping means which grasps said situation, and

10 a synthesizing controller for allowing said voice signal synthesizer to generate said voice signal in which a tone quality including at least one of intonation, volume, speed and key is changed, according to said situation grasped in said grasping means, when said voice signal synthesizer generates 15 said voice signal from said text-based document.

Here, the grasping means grasps a situation in which road guidance is to be outputted, and a situation in which operation guidance is to be outputted. More preferably, it grasps a situation in which VICS 20 information is to be outputted, and a situation in which network information via the Internet is to be outputted.

According to the present invention as described above, even if there are plural types of documents 25 other than narrative in which a hometown and the like of a document creator can be recognized, a voice

intonation and the like are changed according to a length or contents of the text-based document, or a situation what type of voice output is to be made. Therefore, it is possible to give a sense of realism to a listener and even when the document is switched to a different type of document, it can be easily perceived by the listener.

#### **Brief Description of the Drawings**

10 Fig. 1 is a functional block diagram of a navigation system according to the first embodiment of the present invention.

Fig. 2 is flowchart showing an operation of a voice control section according to the first 15 embodiment of the present invention.

Fig. 3 is flowchart showing an operation of the voice control section according to the second embodiment of the present invention.

Fig. 4 is an explanatory diagram showing 20 parameter settings for various types of situations according to the second embodiment of the present invention.

#### **Description of the Preferred Embodiments**

25 Preferred embodiments relating to the present invention will be explained with reference to the

attached drawings.

Referring to Fig. 1 and Fig. 2, a navigation system functioning as a voice output unit of the present invention will be explained.

As shown in Fig. 1, the navigation system 10 of the present embodiment comprises a GPS sensor 11 which receives a signal from a GPS (Global Positioning System) satellite, a VICS information sensor (VICS information receiving means) 12 which receives VICS (Vehicle Information and Communication System) information and a DVD unit 13 which reproduces DVD (Digital Versatile Disc) 1 with map information stored thereon, a communication interface (network information receiving means) 14 which transmits and receives data with a mobile phone 2, a display panel 15, a drive circuit 16 for driving the display panel 15, a speaker 17, a drive circuit 18 for driving the speaker 17, and an operation terminal 19 for various input operations.

Furthermore, this navigation system 10 comprises a route determining section 21 which determines a scheduled route and a guide point, based on a destination inputted by the operation on the operation terminal 19 and a current position obtained from the GPS sensor 11, a guide point detecting section 22 which determines whether or not the current position thus

obtained from the GPS sensor 11 is a guide point, a first text storing section 23 which stores the VICS information obtained from the VICS information sensor 12, and network information such as news, E-mail and the like obtained from a mobile phone 2 via the Internet, 5 a second text storing section 26 which stores a predetermined guidance such as road guidance and operation guidance of the system, a display control section 29 which controls a display output of the display panel 15, and a voice control section 30 which 10 controls a voice output from the speaker 17.

The first text storing section 23 comprises a VICS information text storing section 24 which stores VICS information text, and a network information text 15 storing section 25 which stores network information. The second text storing section 26 comprises a road guidance text storing section 27 which previously stores a road guidance text, and an operation guidance text storing section 28 which previously stores 20 operation guidance text of the system.

The voice control section 30 comprises a grasping section 31, which grasps a situation to know what text is to be vocally outputted in response to a signal from the guide point detecting section 22 or the operation 25 terminal 19, takes out a corresponding text from the storing sections 23 and 26, and recognizes the length

of the text, a voice signal synthesizing section 33 which converts the text into a voice signal, and a synthesizing control section 32 which controls a generation of the voice signal from the voice signal 5 synthesizing section 33.

In the present embodiment, the DVD unit 13 is employed for reproducing the map information. However, if the storage media on which the map information is stored is another type of media, such 10 as CD (Compact Disc) and IC (Integrated Circuit) card, it is a matter of course to employ a reproducing unit conforming to the media, such as CD unit and IC card reader.

Next, operations of the navigation system will 15 be explained.

The route determining section 21 determines a scheduled route, based on a destination inputted by operating the operation terminal 19 and a current position obtained by the GPS sensor 11, and also 20 determines a guide point as a point as to which the road guidance is to be carried out, on the way of the scheduled route. The display control section 29 obtains the scheduled route from the route determining section 21 in response to the operation of the 25 operation terminal 19, and displays the scheduled route on the display panel 15. The display control

section 29 also displays on the display panel 15 a peripheral map of the current position and the scheduled route within the peripheral map, based on the map information from the DVD 1 reproduced by the 5 DVD unit 13 and the current position obtained by the GPS sensor 11.

When the guide point detecting section 22 detects that any one of the plurality of guide points determined by the route determining section 21 becomes 10 the current position indicated by the GPS sensor, a notification is made to the display control section 29 and the voice control section 30. When the display control section 29 receives the notification above, it displays on the display panel 15 a predetermined 15 image to be displayed on the pertinent guide point. For example, when the guide point is positioned at 400 m before a cross point where a right-turn is to be made, an image displayed on the display panel 15 is a detailed map around the cross point, a scheduled route within 20 the detailed map, and so on. When the voice control section 30 receives the above notification, it reads out a road guidance text corresponding to the notification, from the texts stored in the road guidance text storing section 27, converts the road 25 guidance text into a voice signal, and outputs the voice signal from the speaker 17.

When the VICS information sensor 12 receives VICS information, the display control section 29 and the voice control section 30 are notified of the VICS information, and it is stored in the VICS information text storing section 24 of the first text storing section 23. When the display control section 29 receives the above notification, it reads out the VICS information text stored in the VICS information text storing section 24, and displays the text on the display panel 15. When the voice control section 30 receives the above notification, it reads out the VICS information text stored in the VICS information text storing section 24, and converts the VICS information text into a voice signal and outputs the signal from the speaker 17.

When the communication interface 14 receives network information such as E-mail and news from the mobile phone 2, the network information is stored in the network information text storing section 25. When the voice control section 30 receives a voice output notification as to the network information or the operation guidance by operating the operation terminal 19, it reads out a network information text or an operation guidance text in response to the notification, out of the texts stored in the network information text storing section 25 or the texts stored

in the operation guidance text storing section 28, converts the network information text or the operation guidance text into a voice signal and outputs the voice signal from the speaker 17.

5 Next, detailed operations of the voice control section 30 will be explained with reference to the flowchart as shown in Fig. 2.

At first, it is determined whether or not the grasping section 31 of the voice control section 30 10 is in a state of voice outputting (step 1). This determination is made based on whether or not there is any input of a signal from the guide point detecting section 22 or the VICS information sensor 12, or there is any input of a signal instructing a voice output 15 by operating the operation terminal 19. When the grasping section 31 receives a signal from the guide point detecting section 22 or the like, and determines as being in a state of voice outputting, it grasps the current situation to know what kind of voice output 20 is to be made based on the signal (steps 2 to 5). Specifically, it grasps the current situation to know, whether the voice output is to be made regarding VICS 25 information (step 2), network information (step 3), a road guidance (step 4), or an operation guidance (step 5).

Subsequently, the grasping section 31 reads out

a text according to the situation grasped in steps 2 to 5 from the storing sections 23 and 24 (steps 6 to 9), and grasps the length of the text to determine whether or not it is within a predetermined length.

5 Then, the text and the result of the determination are passed to the synthesizing control section 32 (step 10). Here, as the predetermined length of the text, it is set to approximately 100 bytes. As the predetermined length of the text is set to 100 bytes,

10 most of the road guidance texts and operation guidance texts are treated as short. On the other hand, most of the VICS information texts and the network information texts are treated as long.

When the length of the text thus passed is within a predetermined length, that is, it is a short text, the synthesizing control section 32 sets an intonation parameter defining a voice intonation, to a predefined large value, and passes both the intonation parameter and the text to the voice signal synthesizing section 33 (step 11). On the other hand, when the text thus passed is long, the synthesizing control section 32 sets the intonation parameter to a predefined small value, and passes both the intonation parameter and the text to the voice signal synthesizing section 33 (step 12).

The voice signal synthesizing section 33 converts

the text passed from the synthesizing control section 32 into a voice signal. At this stage, the voice signal is generated by use of the intonation parameter passed from the synthesizing control section 32 (step 13).  
5 Here, when a small value is set as the intonation parameter, the voice is made less inflective, and when a large value is set as the intonation parameter, the voice is made more inflective. Therefore, the road guidance or operation guidance constructed by a short  
10 sentence becomes more inflective, and the VICS information or network information constructed by relatively long sentence becomes less inflective. The voice signal synthesizing section 33 outputs thus generated voice signal to the driving circuit 18, and  
15 then outputs the voice from the speaker 17 (step 14).

As described above, since the voice intonation is varied according to the length of the text, it is possible to give the listener a sense of realism, even if there are plural types of documents other than  
20 narrative or the like in which a hometown or the like of a document creator can be recognized. Further, in the present embodiment, since the voice of the road guidance or operation guidance is made more inflective, a driver can be reminded that important information  
25 is now being outputted.

In the embodiment above, only the voice

intonation is changed according to the text length, but also a voice speed, volume or key other than the intonation may be varied simultaneously. Here, a length of the text is grasped, but it may also be 5 possible to grasp contents of the text and to vary the voice intonation and the like according to the contents. As to the text contents, upon reading out a text from each of the storing sections 23 and 26, it is possible to grasp whether the text is a road guidance, network 10 information or the like, by referring to a header part of the text.

Next, with reference to Fig. 3 and Fig. 4, a navigation system according to the second embodiment of the present invention will be explained.

15 The configuration of the navigation system of the present embodiment is basically same as that of the first embodiment as described above with Fig. 1. The Navigation system of the present embodiment, however, is different from the first embodiment in operations 20 of the grasping section 31 and the synthesizing control section 32 of the voice control section 30.

In the following, only the operation of the voice control section 30 of the present embodiment will be explained with reference to Fig. 3.

25 At first, similar to the first embodiment, the grasping section 31 of the voice control section 30

determines whether or not it is a situation to output a voice, according to the existence of a signal from the guide point detecting section 22 or the like (step 1). Then, the grasping section 31 grasps the current 5 situation to know what kind of voice output is to be made, based on the signal from the guide point detecting section 22 or the like (Steps 2 to 5). In other words, as described above, the current situation is grasped to know, whether the voice output is to be 10 made regarding VICS information (step 2), network information (step 3), a road guidance (step 4), or an operation guidance (step 5).

Subsequently, the grasping section 31 reads out a text according to the situation thus grasped in steps 15 2 to 5 (steps 6 to 9), and passes to the synthesizing control section 32 both the text and the situation previously grasped.

The synthesizing control section 32 sets a voice intonation parameter, a voice speed parameter, a voice 20 volume parameter, and a voice key parameter, and passes both these parameters and the text to the voice signal synthesizing section 33 (Steps 20 to 23).

Specifically, each parameter has settings as shown in Fig. 4. As for the VICS information, each parameter 25 is defined such that the intonation is small, the speed and the volume are medium, and the key is high (Step

20). As for the network information, each parameter is defined such that the intonation is small, the speed is high, the volume is small, and the key is high (Step 21). As for the road guidance, each parameter is 5 defined such that the intonation is large, the speed is low, the volume is large and the key is low (step 22). As for the operation guidance, each parameter is defined such that the intonation is large, the speed is low, the volume is large and the key is medium (step 10 23). The settings for each parameter are not limited to those as described above. Further, the settings for each parameter may be defined freely with the operation of the operation terminal 19 by the driver himself or herself, since there are preferences for 15 the settings depending on the driver, i.e., the listener, such as male or female, or a younger person or an elder person.

The voice signal synthesizing section 33 converts the text passed from the synthesizing control section 20 32. At this timing, the voice signal is generated by use of each parameter passed from the synthesizing control section 32 (step 13). Then, the generated voice signal is outputted to the driving circuit 18, and voice is outputted from the speaker 17 (step 14).

25 As described above, according to the present embodiment, it is possible to vary the intonation,

speed and the like, according to the situation what kind of voice is to be outputted.